

[6] —, “A practical method for approaching the channel capacity of constrained channels,” *IEEE Trans. Inform. Theory*, vol. 43, pp. 1389–1399, 1997.

[7] K. J. Kerpez, “Runlength codes from source codes,” *IEEE Trans. Inform. Theory*, vol. 37, pp. 682–687, 1991.

[8] M. Mansuripur, “Enumerative modulation coding with arbitrary constraints and post-modulation error correction coding and data storage systems,” *Proc. SPIE*, vol. 1499, pp. 72–86, 1991.

[9] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*. Amsterdam, The Netherlands: North-Holland, 1977.

[10] B. H. Marcus, “Sofic systems and encoding data,” *IEEE Trans. Inform. Theory*, vol. IT-31, pp. 366–377, 1985.

[11] G. N. N. Martin, G. G. Langdon, and S. J. P. Todd, “Arithmetic codes for constrained channels,” *IBM J. Res. Develop.*, vol. 27, pp. 94–106, 1983.

[12] B. H. Marcus, R. M. Roth, and P. H. Siegel, “Constrained systems and coding for recording channels,” in *Handbook of Coding Theory*, V. S. Pless and W. C. Huffman, Eds. Amsterdam, The Netherlands: Elsevier, 1998, pp. 1635–1764.

[13] B. H. Marcus, P. H. Siegel, and J. K. Wolf, “Finite-state modulation codes for data storage,” *IEEE J. Select. Areas Commun.*, vol. 10, pp. 5–37, 1992.

[14] W. W. Peterson and E. J. Weldon Jr., *Error-Correcting Codes*, 2nd ed. Cambridge, MA: MIT Press, 1972.

[15] R. M. Roth, “On runlength-limited coding with dc control,” *IEEE Trans. Commun.*, vol. 48, pp. 000–000, Mar. 2000.

[16] S. J. P. Todd, G. N. N. Martin, and G. G. Langdon, “A general fixed rate arithmetic coding method for constrained channels,” *IBM J. Res. Develop.*, vol. 27, pp. 107–115, 1983.

$Y_i$	1	0	0	1	0	0	0	1	0
$X_i$	-1	-1	-1	+1	+1	+1	+1	-1	-1

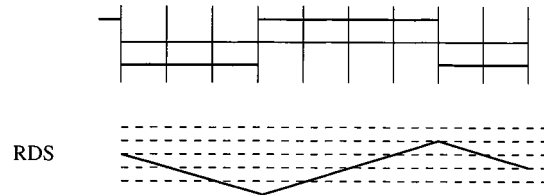


Fig. 1. Running digital sum (RDS) versus time. The binary input symbols  $y_i$  are translated into the bipolar channel bits  $x_i$  using a precoder and a range converter. The other curves show the relationship between the sequence  $\{x_i\}$ , the write signal, and the RDS.

translated into the bipolar channel bits  $x_i$  using a precoder and a range converter. The precoder’s operation can be written as

$$\hat{x}_i = \hat{x}_{i-1} \oplus y_i, \quad x_i = 2\hat{x}_i - 1, \quad 1 \leq i \leq n \quad (1)$$

where  $\oplus$  denotes GF(2) addition. The various signals are shown in Fig. 1.

In this correspondence, we will focus on state-dependent codes of rate  $m/n$ . In its simplest form, the set of encoder states, called *principal states*, is a subset of the  $N$  channel states  $\{z_i\}$ . From each of the principal states there are at least  $2^m$  constrained words beginning at such a state and ending in a principal state. The set of principal states can be found by invoking Franaszek’s procedure [2]. The structure of the codes at hand allows a simpler approach.

Let  $\mathbf{x}$  denote the  $n$ -tuple  $(x_1, \dots, x_n)$ , where  $x_i \in \{-1, 1\}$ , then we define the disparity of  $\mathbf{x}$  as

$$d(\mathbf{x}) = \sum_{i=1}^n x_i.$$

If we require that the running digital sum after concatenation of a new codeword is not larger (in absolute terms) than that at the beginning then each source word must have representation of zero-disparity or it must have two alternative representations of opposite disparity. The words available can easily be computed. Let  $N_-$  and  $N_+$  denote the number of codewords with  $d(\mathbf{x}) \leq 0$  and  $d(\mathbf{x}) \geq 0$ , respectively. Then, we find

$$N_- = N_+ = \begin{cases} 2^{n-1}, & n \text{ odd} \\ 2^{n-1} + \frac{1}{2} \binom{n}{\frac{n}{2}}, & n \text{ even.} \end{cases} \quad (2)$$

When  $n$  is even, we are in the comfortable position that we can choose  $2^{n-1}$  codewords from the many candidates available. For  $n$  odd, on the other hand, it can be seen that there are just enough codewords available to cater to a rate  $(n-1)/n$  code. The implementation of the latter code is usually termed *polarity switch* code, where  $(n-1)$  source symbols are supplemented by one symbol called the *polarity bit* [1], [3]. The encoder has the option to transmit the  $n$ -bit words “as is” or to invert all symbols. The choice of a specific translation is made in such a way that the RDS after transmission of the word is as close to zero as possible. The polarity bit is used by the decoder to undo the action of the encoder. Spectral properties of the polarity bit code have been investigated by Greenstein [4] and Immink [1]. From the above deliberation, it seems, at first glance, that, for  $n$  odd, no codes other than the polarity switch method are feasible. In the next section, we will demonstrate otherwise.

## DC-Free Codes of Rate $(n-1)/n$ , $n$ Odd

Kees A. Schouhamer Immink, *Fellow, IEEE*

**Abstract**—We report on a new class of dc-free codes of rate  $(n-1)/n$ ,  $n$  is odd. Spectral and runlength properties of the new codes have been evaluated by computer simulation.

**Index Terms**—Constrained code, dc-free code, magnetic recording.

### I. INTRODUCTION

Binary sequences with spectral nulls at zero frequency, also called *matched spectral null codes*, have found widespread application in communication and recording systems. Let  $\{x_0, \dots, x_i, \dots\}$ ,  $x_i \in \{-1, 1\}$  be a bipolar sequence. The (running) digital sum  $z_i$  is defined as

$$z_i = \sum_{j=-\infty}^i x_j = z_{i-1} + x_i.$$

If  $z_i$  is bounded, the spectral density of the sequence  $\{x_i\}$  vanishes at zero frequency [1]. Then, at any instant  $i$  the running digital sum (RDS)  $z_i$  assumes a finite number of sum values  $N$ , called *digital sum variation* (DSV). Traditionally, binary input symbols  $y_i \in \{0, 1\}$  are

Manuscript received June 23, 1999; revised October 25, 1999. The material in this paper was presented in part at the IEEE International Symposium on Information Theory, MIT, Cambridge, MA, August 16–21, 1998.

K. A. S. Immink is with the Institute for Experimental Mathematics, 45326 Essen, Germany (e-mail: immink@exp-math.uni-essen.de).

Communicated by E. Soljanin, Associate Editor for Coding Techniques. Publisher Item Identifier S 0018-9448(00)01668-0.

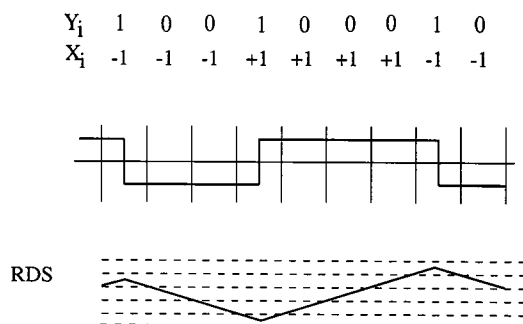


Fig. 2. Running digital sum (RDS) versus time. The binary input symbols  $y_i$  are translated into the bipolar channel bits  $x_i$  using a precoder and a range converter. The transitions occur halfway the bit cells. Note that the contribution to the RDS of the 9-bit word  $\mathbf{y}$  is nil, while, see Fig. 1, using the conventional definition the contribution of the same input word  $\mathbf{y}$  is  $-1$ .

TABLE I  
SUM VARIANCE OF THE NEW CODE  $s^2$   
AND THE SUM VARIANCE  $s_p^2$  OF THE POLARITY SWITCH CODE VERSUS  
CODEWORD LENGTH  $n$

$n$	$s^2$	$s_p^2$
3	1.05	1.67
5	1.73	3.00
7	2.52	4.33
9	3.23	5.66
11	4.07	7.00

## II. THE NEW CONSTRUCTION

The crux of the new class of dc-free code is the redefinition of the running digital sum. In the classical definition, see Fig. 1, it is assumed that transitions occur at the start of the bit cells. In the new definition, transitions are assumed to occur halfway the bit cells. Fig. 2 shows the various signals. Note that the new definition does not have a physical meaning. Essentially, the write signal is delayed by half a channel bit interval. Let  $x_0$  be the value of the bit preceding the word  $\mathbf{x}$ . Then the disparity,  $ds(\mathbf{y}, x_0)$ , of the binary  $n$ -tuple  $\mathbf{y}$  after precoding using the new definition is

$$ds(\mathbf{y}, x_0) = \sum_{i=1}^n x_i + (w(\mathbf{y}) \bmod 2)x_0 \quad (3)$$

where  $w(\mathbf{y})$  denotes the weight of  $\mathbf{y}$ . The first right-hand term is the “conventional” contribution to the RDS, and the second term expresses a correction term. The interesting result of this new definition is that zero-disparity, i.e.,  $ds(\mathbf{y}, x_0) = 0$ , codewords are possible for  $n$  odd.

Let  $N(s)$  denote the number of words  $\mathbf{y}$  having  $ds(\mathbf{y}, 1) = s$ . Then, it is not difficult to see that  $N(s)$  can be found with

$$N(s) = \binom{n-1}{\frac{n-1}{2} + \lfloor \frac{s}{2} \rfloor} \quad (4)$$

The zero-disparity words are uniquely allocated to the source words. The other codewords are allocated in pairs of opposite disparity. The choice of a specific representation is made to minimize the absolute value of the running digital sum. The words available in both modes can easily be computed. To that end, let  $\mathcal{N}_-$  and  $\mathcal{N}_+$  denote the sets whose

members  $\mathbf{y}$  satisfy  $ds(\mathbf{y}, 1) \leq 0$  and  $ds(\mathbf{y}, 1) \geq 0$ , respectively, and let  $N_-$  and  $N_+$  denote the cardinality of these sets. Then

$$N_- = 2^{n-1} \quad (5)$$

and

$$N_+ = 2^{n-1} + \binom{n-1}{\frac{n-1}{2}} \quad (6)$$

From (5) and (6) we infer that, as  $N_- = 2^{n-1}$ , the designer has no other choice than taking the words available, whereas, as  $N_+ > 2^{n-1}$ , the designer has the freedom to choose words that satisfy certain design criteria such as low coder/decoder complexity, minimizing maximum run length, etc. It can easily be verified that both  $10^{n-1} \in \mathcal{N}_-$  and  $0^{(n-1)/2}10^{(n-1)/2} \in \mathcal{N}_-$ , where  $0^p$  denotes a string of  $p$  zeros. We conclude, therefore, that the maximum “zero”-run length equals  $3(n-3)/2$ .

## III. SPECTRAL PERFORMANCE

The spectral performance of the new codes has been investigated by computer simulation. As an example we studied codes of rate  $8/9$ . The number of codewords with nonnegative disparity,  $N_+$ , equals 326. We require 256 words, and from the codewords available we chose the words with the smallest absolute disparity. The maximum run length is 12. The rate  $8/9$ , polarity-switch code shows a 4-dB less rejection at the low-frequency end. Also the maximum “zero”-run length is much larger, 18, than that of the new code. The quantity *sum variance*, which is inversely proportional to the width of the spectral notch at zero frequency, has been adopted to quantify the low-frequency properties of a dc-free code [1]. Table I lists the sum variance of the new code  $s^2$  and the sum variance  $s_p^2$  of the polarity switch code versus codeword length  $n$ . It can be seen that the new codes perform better as they show a smaller sum variance than that of the conventional scheme.

## IV. CONCLUSIONS

We have reported on a new class of rate  $(n-1)/n$ ,  $n$  odd, dc-free codes which have been designed on the basis of a redefinition of the running digital sum. Properties of the new codes, such as spectral and run-length distribution have been evaluated by computer simulation.

## REFERENCES

- [1] K. A. S. Immink, *Codes for Mass Data Storage Systems*. Eindhoven, The Netherlands: Shannon Foundation Publishers, 1999.
- [2] P. A. Franaszek, “Sequence-state encoding for digital transmission,” *Bell Syst. Tech. J.*, vol. 47, pp. 143–157, Jan. 1968.
- [3] K. W. Cattermole, “Principles of digital line coding,” *Int. J. Electron.*, vol. 55, pp. 3–33, July 1983.
- [4] L. J. Greenstein, “Spectrum of a binary signal block coded for dc suppression,” *Bell Syst. Tech. J.*, vol. 53, pp. 1103–1126, July 1974.